

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-35069

(43) 公開日 平成9年(1997)2月7日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 T 7/00		9061-5H	G 0 6 F 15/70	3 3 0 B
H 0 4 N 1/41			H 0 4 N 1/41	B
7/24			7/13	Z

審査請求 未請求 請求項の数27 O L (全 16 頁)

(21) 出願番号 特願平8-180273

(22) 出願日 平成8年(1996)7月10日

(31) 優先権主張番号 5 0 0 6 7 2

(32) 優先日 1995年7月10日

(33) 優先権主張国 米国 (U S)

(71) 出願人 595119464

エイ・ティ・アンド・ティ・アイピーエム・コーポレーション

アメリカ合衆国, 33134 フロリダ, コーラル ゲーブルズ, ボンス ド レオン
ブウルヴァード 2333(72) 発明者 アレクサンドロス エレフテリアディス
アメリカ合衆国, 10027 ニューヨーク,
ニューヨーク, モーニングサイド ドライ
ブ 100, アパートメント4-イー

(74) 代理人 弁理士 三俣 弘文

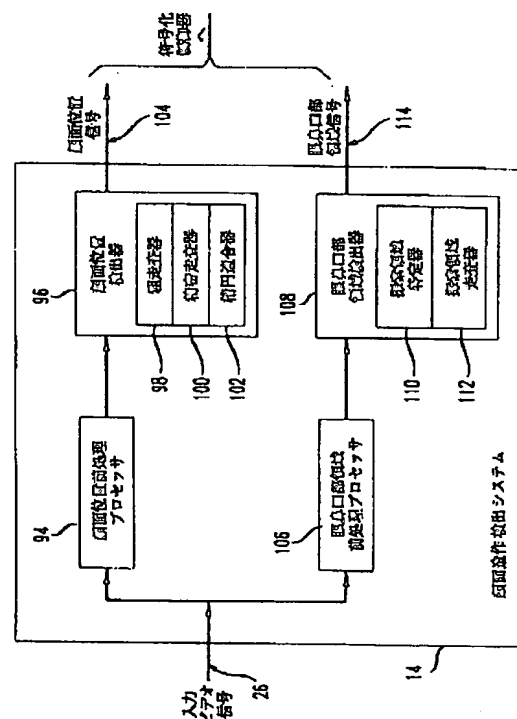
最終頁に続く

(54) 【発明の名称】 ビデオ信号符号化装置及び方法

(57) 【要約】

【課題】 ビデオ信号符号化のため被写対象物を自動検出する。

【解決手段】 顔面造作検出システム14が、フレームの連続であってこれらのフレームのうちの少なくとも1個のフレームが対象物の画像に対応するようなフレームの連続を表すビデオ信号26に応動し、対象物の少なくとも1個の領域を検出する。顔面造作検出システム14は、ビデオ信号を処理することにより、少なくとも、対象物の領域のうちの閉曲線の少なくとも一部によって特徴付けられるような対象物の領域、を検出し且つビデオ信号の符号化に用いるために閉曲線に関連する複数のパラメータを生成するためのプロセッサ94、96、106、108からなる。



【特許請求の範囲】

【請求項1】 フレームの連続であって該フレームのうちの少なくとも1個のフレームが対象物の画像に対応するようなフレームの連続を表すビデオ信号を符号化するための、ビデオ信号符号化装置であって、

前記ビデオ信号を処理することにより、少なくとも、前記対象物の領域のうちの閉曲線の少なくとも一部によって特徴付けられるような前記対象物の領域、を検出し且つ前記ビデオ信号の符号化に用いるために前記閉曲線に関連する複数のパラメータを生成するためのプロセッサからなることを特徴とするビデオ信号符号化装置。

【請求項2】 前記プロセッサが、少なくとも、前記対象物の領域のうちの前記閉曲線としての楕円の少なくとも一部によって特徴付けられるような前記対象物の領域、を検出するようにしたことを特徴とする請求項1の装置。

【請求項3】 前記プロセッサが、少なくとも、前記対象物としての頭部外形の領域のうちの該頭部外形にほぼ適合する前記閉曲線としての前記楕円の少なくとも一部によって特徴付けられるような前記対象物としての頭部外形の領域を検出するようにしたことを特徴とする請求項2の装置。

【請求項4】 前記プロセッサが、少なくとも、前記対象物の領域のうちの前記閉曲線としての矩形の少なくとも一部によって特徴付けられるような前記対象物の領域を検出するようにしたことを特徴とする請求項1の装置。

【請求項5】 前記プロセッサが、少なくとも、前記対象物としての頭部の眼部領域のうちの該眼部領域の対称軸にほぼ平行な対称軸を有する前記閉曲線としての前記矩形の少なくとも前記一部によって特徴付けられるような前記対象物としての頭部の眼部領域を検出するようにしたことを特徴とする請求項4の装置。

【請求項6】 前記プロセッサが、少なくとも、前記対象物の領域のうちの前記閉曲線としての台形の少なくとも一部によって特徴付けられるような前記対象物の領域を検出するようにしたことを特徴とする請求項1の装置。

【請求項7】 前記プロセッサが、少なくとも、前記対象物としての頭部の眼鼻口部領域のうちの該眼鼻口部領域にほぼ適合する前記閉曲線としての前記台形の少なくとも前記一部によって特徴付けられるような前記対象物としての頭部の眼鼻口部領域を検出するようにしたことを特徴とする請求項6の装置。

【請求項8】 前記プロセッサが、前記ビデオ信号を処理することにより、前記対象物の前記領域の縁部に対応する縁部データ信号を生成するための前処理プロセッサと、

前記縁部データ信号を処理することにより前記複数のパラメータを生成するための対象物検出器とをさらに有す

ることを特徴とする請求項1の装置。

【請求項9】 前記パラメータに関する対象物検出信号に応動して、前記対象物の前記検出された領域の符号化を制御する目的でバッファサイズを増加させるため前記ビデオ信号の符号化に用いられる量子化器ステップサイズを調整するためにバッファサイズ変調を行うための符号化制御器をさらに有することを特徴とする請求項1の装置。

【請求項10】 前記パラメータに関する対象物検出信号に応動して、前記対象物の前記検出された領域の符号化を制御する目的で符号化のレートを増加させるため前記ビデオ信号の符号化に用いられる量子化器ステップサイズを調整するためにバッファレート変調を行うための符号化制御器をさらに有することを特徴とする請求項1の装置。

【請求項11】 対象物の検出された領域を表示する対象物検出信号に応動して、フレームの連続であって該フレームのうちの少なくとも1個のフレームが対象物の画像に対応するようなフレームの連続を表すビデオ信号の符号化を制御するための、ビデオ信号符号化制御装置であって、

バッファサイズ変調を行うための、且つ前記対象物の前記検出された領域を符号化する目的でバッファサイズを増加させるために前記ビデオ信号の符号化に用いられる量子化器ステップサイズを調整するためのプロセッサからなることを特徴とするビデオ信号符号化制御装置。

【請求項12】 前記プロセッサが、CCITT勧告H.261の標準に従って符号化する目的で前記量子化器ステップサイズを調整するために前記バッファサイズ変調を行うようにしたことを特徴とする請求項11の装置。

【請求項13】 対象物の検出された領域を表示する対象物検出信号に応動して、フレームの連続であって該フレームのうちの少なくとも1個のフレームが対象物の画像に対応するようなフレームの連続を表すビデオ信号の符号化を制御するための、ビデオ信号符号化制御装置であって、

バッファレート変調を行うための、且つ前記対象物の前記検出された領域の符号化のレートを増加させる目的で前記ビデオ信号の符号化に用いられる量子化器ステップサイズを調整するためのプロセッサからなることを特徴とするビデオ信号符号化制御装置。

【請求項14】 前記プロセッサが、CCITT勧告H.261の標準に従って符号化する目的で前記量子化器ステップサイズを調整するために前記バッファレート変調を行うようにしたことを特徴とする請求項13の装置。

【請求項15】 フレームの連続であって該フレームのうちの少なくとも1個のフレームが対象物の画像に対応するようなフレームの連続を表すビデオ信号に応動して

対象物の領域を符号化するためのビデオ信号符号化方法であって、

前記対象物の領域のうちの閉曲線の少なくとも一部によって特徴付けられるような前記対象物の領域を検出するステップと、

前記閉曲線に連関する複数のパラメータを生成するステップと、

前記複数のパラメータを用いて前記ビデオ信号を符号化するステップとからなることを特徴とするビデオ信号符号化方法

【請求項16】 前記検出するステップが、少なくとも、前記対象物の領域のうちの前記閉曲線としての楕円の少なくとも一部によって特徴付けられるような前記対象物の領域を検出するステップからなるようにしたことを特徴とする請求項15の方法。

【請求項17】 前記検出するステップが、少なくとも、前記対象物としての頭部外形の前記領域を検出するステップと、前記頭部外形が前記閉曲線としての楕円の少なくとも一部によって特徴付けられるように、該楕円を前記頭部外形にほぼ適合させるステップとからなるようにしたことを特徴とする請求項16の方法。

【請求項18】 前記検出するステップが、少なくとも、前記対象物の領域のうちの前記閉曲線としての矩形の少なくとも一部によって特徴付けられるような前記対象物の領域を検出するステップからなるようにしたことを特徴とする請求項15の方法。

【請求項19】 前記検出するステップが、少なくとも、前記対象物としての頭部の眼部領域を検出するステップと、前記眼部領域の対称軸にほぼ平行な対称軸を有するような前記閉曲線としての前記矩形の少なくとも一部、を定めるステップとからなるようにしたことを特徴とする請求項18の方法。

【請求項20】 前記検出するステップが、少なくとも、前記対象物の領域のうちの前記閉曲線としての台形の少なくとも一部によって特徴付けられるような前記対象物の領域を検出するステップからなるようにしたことを特徴とする請求項15の方法。

【請求項21】 前記検出するステップが、少なくとも、前記対象物としての頭部の眼鼻口部領域を検出するステップと、前記頭部の前記眼鼻口部輪郭が前記閉曲線としての台形の少なくとも一部によって特徴付けられるように、該台形を前記頭部外形にほぼ適合させるステップとからなるようにしたことを特徴とする請求項20の方法。

【請求項22】 前記複数のパラメータを生成するステップが、前記ビデオ信号を前処理することにより、前記対象物の前記領域の縁部に対応する縁部データ信号を生成するた

めの前処理ステップからなり、

前記検出するステップが、

前記縁部データ信号を処理することにより、閉曲線の少なくとも一部によって特徴付けられるような前記対象物の少なくとも前記領域を検出するステップからなるようにしたことを特徴とする請求項15の方法。

【請求項23】 前記対象物検出信号に応動して、量子化器ステップサイズを調整するステップと、前記調整された量子化器ステップサイズを用いて前記ビデオ信号を符号化するステップとをさらに有することを特徴とする請求項15の方法。

【請求項24】 前記対象物の前記検出された領域を符号化するためにバッファサイズを増加させるステップをさらに有することを特徴とする請求項23の方法。

【請求項25】 前記対象物の前記検出された領域を符号化するレートを増加させるステップをさらに有することを特徴とする請求項23の方法。

【請求項26】 対象物の検出された領域を表示する対象物検出信号に応動して、フレームの連続であって該フレームのうちの少なくとも1個のフレームが該対象物の画像に対応するようなフレームの連続を表すビデオ信号の符号化を制御するための、ビデオ信号符号化制御方法であって、

前記対象物検出信号に応動して、量子化器ステップサイズを調整するステップと、前記調整された量子化器ステップサイズを用いて、前記ビデオ信号を符号化するステップとからなることを特徴とするビデオ信号符号化制御方法。

【請求項27】 前記対象物の前記検出された領域を符号化するためにバッファサイズを増加させるステップをさらに有することを特徴とする請求項26の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、概して信号符号化の分野に関し、詳しくは映像（ビデオ）の符号化に用いられる顔面造作の検出に関する。

【0002】

【従来の技術】ラップトップ型等を含むパーソナルコンピュータ（PC）における計算力の増加につれて、テキスト、音響、及びビデオの機能を組み込んだマルチメディアのアプリケーションがますますPCに利用可能になって来ている。

【0003】動画処理（MPEG）標準の場合のように、ビデオは概して、高品質の画像（イメージ）を得るのに比較的高いビットレートを要するが、テレビ会議のような場合には、比較的低いビットレートをを用いることが許容される。

【0004】

【発明が解決しようとする課題】このような低いビットレートのテレビ会議においては、ビデオの符号化によ

10

20

30

40

50

5

て、符号化された画像全体に系統的に、人工的な部分が生じる。これは例えば、画像のうちの顔面部分と非顔面部分との両方を同じビットレートで符号化することに起因する。

【0005】画像を見る人が目を画像中の人々の目に合わせること（アイコンタクト）（目合わせ）を継続すること等によって顔面の造作に焦点を合わせようとする、画像の顔面部分と非顔面部分との両方について符号化を共通のビットレートで行うことになるため、顔面部分の符号化に十分な品質を得ることができず、その結果としてビデオ画像が見えにくくなる。そして、読唇に頼る聴覚障害者の場合のような或る種の状況下では、顔面の造作の非常に良好な表現が理解度にとって至上的重要性を有することがある。

【0006】

【課題を解決するための手段】本発明に基づく装置は、フレームの連続であつこれらのフレームのうちの少なくとも1個のフレームが対象物の画像に対応するようなフレームの連続を表すビデオ信号に応動し、対象物の少なくとも1個の領域を検出する。

【0007】本発明に基づく装置は、ビデオ信号を処理することにより、少なくとも、対象物の領域のうちの閉曲線の少なくとも一部によって特徴付けられるような対象物の領域、を検出し且つビデオ信号の符号化に用いるために閉曲線に関連する複数のパラメータ（媒介変数）を生成するためのプロセッサからなる。

【0008】

【発明の実施の形態】以下、図面を詳細に参照して本発明に基づき、ビデオにおける顔面造作の検出及び位置探索のための顔面造作検出システムについて説明する。図中、同様な符号は類似又は同一の要素を示すものとする。

【0009】本発明の顔面造作検出システム及び方法は、一実施例において、図1の低ビットレートのビデオ「符号化復号化器（コーデック）」10において実現される。この実現は、 $p \times 64$ キロビット/秒（k b p s）（ここに $p = 1, 2, \dots, 30$ ）のビットレートでのビデオ符号化のためのCCITT勧告H.261に適合したビデオ符号化方法のような、動作補償離散余弦変換（MC-DCT）手法に基づいて行われる。

【0010】図1に示す実施例において、コーデック10は、マイクロプロセッサのようなプロセッサとメモリ（図示しない）とを有する対象物位置探索器12からなる。この対象物位置探索器12は、対象物検出プログラムを実行して、本発明の顔面造作検出システム14及び方法を実現する。一実施例において、対象物検出プログラムは、Cプログラミング言語で書かれた、コンパイルされたソースコードである。

【0011】コーデック10は又、マイクロプロセッサのようなプロセッサとメモリ（図示しない）とを有する

6

符号化制御器16からなる。この符号化制御器16は、符号化制御プログラムを実行して、本発明のバッファレート変調器18及びバッファサイズ変調器20を実現する。一実施例において、符号化制御プログラムは、C++プログラミング言語で書かれた、コンパイルされたソースコードである。

【0012】なお、本発明の顔面造作検出システム14及び方法を含む対象物位置探索器12並びに本発明のバッファレート変調器18及びバッファサイズ変調器20を含む本発明の符号化制御器16は、下で述べる動作を行うように、他のプログラミング言語を用いてソフトウェア的に、そして/又はハードウェア又はファームウェアによって、実現されることを理解されたい。

【0013】説明を分かりやすくするため、本発明の顔面造作検出システム14及び方法を含む対象物位置探索器12並びに本発明のバッファレート変調器18及びバッファサイズ変調器20を含む本発明の符号化制御器16は、「プロセッサ」と名付けた機能ブロックを含む個々の機能ブロックを有するものとして説明する。これらのブロックによって表される機能は、ソフトウェアを実行する能力を有するハードウェアを含むがこれに限らない共用又は専用のハードウェアを用いることによって得られる。

【0014】例えば、ここに示すブロックの機能は、単一の共用プロセッサによっても、又は複数の個々のプロセッサによっても得られる。又、図中の符号をつけた機能ブロックを用いていても、それがソフトウェアの実行能力を有するハードウェアを専用的に指すと解釈すべきではない。

【0015】図示の実施例は、AT&T社製DSP16又はDSP32Cのようなデジタル信号処理装置（DSP）ハードウェア、下で述べる動作を行うソフトウェアを格納するための読み出し専用メモリ（ROM）、及びデジタル信号処理装置の処理結果を格納するための読み出し専用メモリ（ROM）を有する。超大規模集積回路（VLSI）ハードウェア実施例及び汎用DSP回路と組み合わせたカスタムVLSI回路も設けられる。これらの実施例は全て、ここに用いる機能ブロックの範囲に入る。

【0016】コーデック10は更に又、符号化制御器16によって制御されるビデオ符号器22及びビデオ復号器24を有する。符号化動作に関して、コーデック10が、入力ビデオ信号26及び外部制御信号28を受信し、対象物位置探索器12及びビデオ符号器22が外部制御信号28を用い符号化制御器16からの制御の下に符号化ビデオ信号26を処理して、符号化されたビットストリーム（出力）30を生成する。

【0017】一実施例において、ビデオ符号器22が、ソース符号器32、ビデオ多重（マルチプレックス）符号器34、送信バッファ36、及び送信符号器38を用

7

いて入力ビデオ信号26を処理し、符号化されたビットストリーム（出力）30を生成する。

【0018】復号化動作に関して、コーデック10が、符号化されたビットストリーム（入力）40をビデオ復号器24によって受信し、ビデオ復号器24が、受信復号器42、受信バッファ44、ビデオ多重（マルチプレックス）復号器46、及びソース復号器48を用いて符号化されたビットストリーム（入力）40を処理して、出力ビデオ信号50を生成する。

【0019】モデルに補助された符号化を用いるビデオ符号器の例が、米国特許出願第08/250,251号（1994年5月27日出願、名称“MODEL-ASSISTED CODING OF VIDEO SEQUENCES AT LOW BIT RATES”）に詳しく述べられている。

【0020】図1のビデオ符号器22の、モデルに補助されたソース符号器32の実現例を図2に詳しく示す。図2において、ソース符号器32が、量子化された変換係数信号（「q」と名付ける）、動きベクトル信号（「v」と名付ける）、及びループフィルタ68のオンオフ切替を表示する切替信号（「f」と名付ける）、のような出力信号を生成するための構成要素54〜70を有する。

【0021】符号化制御器16が、対象物位置探索器12からの対象物位置信号を用いて、INTRA/INTERフラッグ信号（「p」と名付ける）、送信状態を表示するためのフラッグ（「t」と名付ける）、及び量子化器表示信号（「qz」と名付ける）、のような制御信号を生成する。これらの制御信号は、ビデオ多重符号器34に供給されて更に処理される。

【0022】図1の例示実施例に示すコーデック10、及び例示実施例において図2に示すような、モデルに補助されたソース符号器を有するビデオ符号器22は、文献（“Line Transmission on Non-Telephone Signals - Video Codec For Audiovisual Services at p x 64 kbit/s - Recommendation H.261”, CCITT, Geneva, 1990）に述べられているようなCCITT勧告H.261に基づいて作動する。

【0023】それから、本発明による顔面造作検出システム14及び方法が、下で更に詳しく述べるように実現される。その実現の際、符号化制御器16に供給される対象物位置探索器12からの対象物位置信号が、顔面造作検出システム14からの顔面造作検出信号を有する。

【0024】顔面造作検出システム14及び方法が、テレビ会議のビデオシーケンスを領域に基づいて選択的に符号化するために、顔面区域位置探索情報を自動的に抽出する。それから、複雑度の低い（計算量の少ない）方法が実行されて頭部外形が検出される。そして、ダウンサンプリングされた、バイナリーのしきい値処理された縁部画像から、眼鼻口部領域が識別特定される。符号化

8

一）シーケンスの特性及び内容に関しての制約は存在しない。

【0025】頭部の顕著な回転及び／又は対象物を移動させることによって部分的にデータが閉塞妨害される現象がある場合、並びに画像内の人物の顔面に毛髪が存在しそして／又は画像内の人物が眼鏡をかけている場合に、頭部外形検出及び眼鼻口部検出の動作は、正確にそして確固として機能する。

【0026】本発明は又、図2に示すH.261対応（コンパチブル）のソース符号器32における量子化器58のような標準の符号化システムにおける量子化器を対象物に合わせて選択的に制御する機能を有する。本実施例において、コーデック10は、文献（“Description of Reference Model 8 (RM8)”, CCITT SGXV WG4, Specialists Group on Coding for Visual Telephony, Doc. 525, June 1989）に述べられているような、H.261に従う符号器を参照モデル8（RM8）について実現した符号器である。

【0027】なお、本発明の顔面造作検出システム14及び方法は、H.261に限定されるものではなく、CCITT勧告H.263及びMPEGのような他の別個の符号化手法、及びニューラルネットワークに基づく顔面造作分類器のような他の別個の顔面造作検出装置でも使用できることを理解されたい。

【0028】本発明の顔面造作検出システム14及び方法を含む対象物位置探索器12は、バッファレート変調及びバッファサイズ変調を行うことによって量子化器58を制御する動作を行う。

【0029】すなわち、量子化器58に関連する符号化制御器16のレート制御器に強制制御して非顔面画像区域の符号化から顔面画像区域の符号化へ全利用可能ビットレートの比較的小部分を転送させるようにすることによって制御する。

【0030】例えば平均ビットレート約10〜15%を転送することによって、本発明の顔面造作検出システム14及び方法から、よりよく表現された顔面造作が産出される。例えば、顔面区域におけるブロック状の人工部分の表現明白度が下り、目合わせが保持される。

【0031】本発明の面造作検出システム14及び方法によれば、入力カラービデオ信号用にそれぞれ留保された56kbps及び24kbpsについて、64kbps及び32kbpsのレートで符号化されたビデオシーケンスに知覚的に顕著な改善が得られる。

【0032】コーデック10は、64kbpsの視聴覚総合ISDN（サービス総合デジタルネットワーク）レートを用いて、YUVフォーマットの入力デジタルカラービデオ信号で、そして「ヘッド・アンド・ショルダー」シーケンスを表すビデオ信号について56kbpsの符号化レートで、作動する。本発明の顔面造作検出器及び方法においては、画像内の顔面位置の検出動作及

び「眼鼻口部」領域の検出動作が行われる。

【0033】本実施例において、顔面造作検出システム14が、顔面の側面図又は顔面が画像に関して傾けられている場合の斜視図を含めた顔面を楕円として定める。

「眼鼻口部」領域の検出には、回転させた頭部についての2次元投影に現れる人間の顔面が本質的に有する傾斜顔面軸に関するこれら領域の対称性が用いられる。

【0034】下で述べるように、本実施例において、入力ビデオ信号26は、ビデオデータ及び「ヘッド・アンド・ショルダー」シーケンスを表す。なお、本発明の顔面造作検出システム14及び方法が、他のレートで作動する他のビデオ符号化システム及び手法にも用いられること、又このような対象物追跡（トラッキング）手法が、顔面以外の対象物が関心対象となるような他の用途にも適応できることを理解されたい。

【0035】用語「顔面位置」は本説明においては、「頭部外形」の位置を包含するように、プロフィールにあるような、頭部を左又は右に回した人を表す画像、及びカメラに背を向けた人の画像を含むものとして定義される。

$$ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0, \quad b^2 - ac < 0 \quad (1)$$

ここに、判別式 $D = b^2$ が負の値を取る場合、式(1)によって楕円が定まる。式(1)のパラメータは、下で更に詳しく述べるようにコーデック10に供給される。

【0040】いくつかの場合には、楕円状の頭部外形によって、顔面位置の概略推定が得られる。しかし、楕円状の頭部外形の使用を適応させて、図4に示すような矩形領域82を特定することが、本発明の顔面造作検出システム14及び方法によって可能となる。矩形領域82は、楕円72の中心と共通の中心を座標 (x_0, y_0) に有し、楕円72の長軸に平行する顔面外形の対称軸84 (A_s と名付ける) を有する。

【0041】矩形領域82には、人の眼、鼻及び口が符号化された画像の形で含まれる。このような矩形領域82は、図4の台形領域(Rと名付ける)を用いた顔面外形をモデル化する際に、図4に示すような対称軸84に平行な画像垂直線86に関してその側線の傾斜を許すことにより、例えば図3の楕円72のような楕円だけを用いる場合に比してより高い自由度が得られ、これによって頭部が僅かに動いた場合に確固たる検出が確実に得られるようになる。

【0042】一実施例において、図4中にウィンドウ88 (W_u と名付ける) として定義される矩形82の上部約1/3の部分には、概して人の顔面内で対称性の最も確かな2個の顔面造作である眼部及び眉毛部が含まれる。ウィンドウ88は、座標 (x_i, y_i) を有する中心90、ウィンドウ幅 W_w 、高さ h 、及び台形側線の角度 θ_i によって特徴付けられる。この場合、台形の形状を定める角度92は、概して人の顔面の眼、鼻及び口の相対的位置を表示する。

【0036】顔面位置は、図3に示すようなEと名付ける楕円72によって表される。楕円72は、座標 (x_0, y_0) に位置する中心74、楕円の長軸及び短軸にそれぞれ沿った半長径76 (Aと名付ける) 及び半短径78 (Bと名付ける)、及び連関する傾斜角80 (θ_0 と名付ける) を有する。

【0037】楕円72の長径に沿って対向する両端部の区域が、顔面の上部及び下部区域がそれぞれ位置する部分である。人の毛髪領域及び頸部領域のような実際の顔面72の上部及び下部区域は湾曲状態が全く異なるが、楕円では、顔面外形のモデルとして比較的精度が得られ、又パラメータに関しては単純になる。

【0038】もしこの顔面外形情報が顔面外形の再生成に用いられなかった場合、本発明の顔面造作検出システム及び方法におけるモデル適合精度が比較的欠如するため、符号化プロセス全体としての性能に対して顕著なインパクトが得られないことになる。

【0039】任意のサイズ及び傾斜の楕円は、下の式(1)の、間接形式での二次、ノンパラメトリック方程式

20 式によって表される。

【0043】図5に図示するように、図1～図2の顔面造作検出システム14は、検出器によって顔面造作を検出するための入力ビデオ信号の前処理用に少なくとも1個の前処理プロセッサを有する。図5に示すように、顔面造作検出システム14は、顔面位置検出器96に接続されて作動する顔面位置前処理プロセッサ94を有し、顔面位置検出器96は、粗走査器(スキャナ)98と、精密走査器100と、楕円適合器102とを有して、顔面位置信号104を生成する。

【0044】顔面造作検出システム14は更に、眼鼻口部領域検出器108に接続されて作動する眼鼻口部領域前処理プロセッサ106を有し、眼鼻口部領域検出器108は、探索領域特定器110と探索領域走査器112とを有して、眼鼻口部領域信号114を生成する。探索領域特定器110及び探索領域走査器112の各々は、顔面造作検出システム14に関連して上記したようなハードウェア及び/又はソフトウェアとして実現される。

【0045】各前処理プロセッサ94、106は、図6に示すような前処理回路116として実現される。前処理回路116は、時間ダウンサンプリング器118、低域フィルタ120、デシメータ(減数器)122、縁部検出器124、及びしきい値回路126を用いる。時間ダウンサンプリング器118は、もし入力ビデオ信号26が例えば図2のソース符号器32の、望まれる入力フレームレートまでダウンサンプリングされていない場合に、設けられる。

【0046】時間ダウンサンプリング器118は、例えば約30フレーム/秒(fps)から、ビデオコーデック10への入力ビデオ信号のフレームレートとなる例え

ば約7.5fpsへの、入力輝度ビデオ信号の時間ダウンサンプリング（時間的減数サンプリング）を行う。低域フィルタ120は、カットオフ周波数を π/c （ c はデシメーションファクタ）とする、サイズ 360×240 画素の入力ビデオ信号の空間低域フィルタ処理を行うための分離可能なフィルタである。

【0047】本実施例において、図5に示す顔面位置前処理プロセッサ94及び眼鼻口部領域前処理プロセッサ106の各々は、一般的な時間ダウンサンプリング器及び一般的な低域フィルタを用いるように実現される。低域フィルタ処理後、フィルタ処理された入力ビデオ信号は、デシメータ122によって処理されて、予め定められたサイズの低域パス画像を生成するための、水平及び垂直直次元での予め定められたデシメーションファクタによるデシメーション（減数）が行われる。

$$\delta_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \delta_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2)$$

これらの演算子は、画像傾斜の成分を定めるのに用いられる。

【0050】それから、縁部検出器124を用いて各画素において傾斜のマグニチュードを生成することによって、傾斜マグニチュード画像が得られる。そして、傾斜マグニチュード画像のしきい値処理を行うためにしきい値回路126を用いて、バイナリの縁部データ信号が生成される。

【0051】図5に示す顔面位置検出器96及び眼鼻口部領域検出器108の各々が、それぞれの前処理プロセッサ94、106からのそれぞれのバイナリの縁部データ信号を用いて、入力ビデオ信号26によって表される画像の顔面位置及び眼鼻口部領域をそれぞれ検出する。

【0052】顔面位置検出器96はサイズ 45×30 画素の前処理されしきい値処理された傾斜マグニチュード画像を用いて、楕円として幾何学的にモデル化された顔面位置の輪郭を検出しトレースして、長円形状（すなわち充填された形状）及びデータによって部分的に閉塞妨害された長円輪郭の両方について位置探索を行う。

【0053】顔面位置検出器96は、階層的な3ステップの手順を実行する。すなわち、粗走査器98による粗走査ステップと、精密走査器100による精密走査ステップと、楕円適合器102による楕円適合ステップとの3ステップで、それらの各々は、顔面造作検出システム14について上記したようなハードウェア及び／又はソフトウェアの形で実現される。それから顔面位置検出器96は、多数の候補のうちから最もそうでありそうな顔面外形として、画像内の検出された楕円を選択する。

【0054】これら3ステップによる、認識及び検出タスクの伸長（デコンポジション）が、入力画像サイズが小さいこととあいまって、本発明の顔面造作検出システ

【0048】本実施例において、顔面位置前処理プロセッサ94は、サイズ 45×30 画素の画像を生成するためにファクタ $c=8$ によるデシメーションを行うデシメータを用いる。眼鼻口部領域前処理プロセッサ106は、ダウンサンプリングにおいて例えば眼、鼻、及び口縁部データのような関心対象の顔面造作データを失うことのないように、サイズ 180×120 画素の画像を生成するためにファクタ $c=2$ によるデシメーションを行うデシメータを用いる。

【0049】各前処理プロセッサ94、106はそれから、ソベル演算子手法を用いて縁部検出器124により、デシメーションされた画像について縁部検出処理を行う。ここに、ソベル演算子は、例えば次式のように、水平及び垂直演算子によって行列式で表される。

【数1】

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (2)$$

ム14の計算量を少なくする効果をもたらし、大きな候補者プールを対象とする消耗的な探索が回避される。

【0055】粗走査については、粗走査器98が、入力バイナリ縁部データ信号を、サイズ $B \times B$ 画素、例えばサイズ 5×5 、のブロックに分割（セグメント化）する。各ブロックは、もしそのブロック内の画素のうちの少なくとも1個の画素が非ゼロ値を有する場合には粗走査器98によってマーク付けされる。それからそのブロックの配列が、例えば左から右、上から下という仕方で走査され、マーク付けされたブロックの隣接ランについて探索が行われる。これらの各ランについて、精密走査及び楕円適合の動作が行われる。

【0056】精密走査器100が1個のランのブロック内の画素を例えば左から右、上から下という仕方で走査して、画素の隣接ランに対向する、非ゼロ画素値を有する最初のラインを検出する。検出されたラインの最初及び最後の画素（座標（ X_{START} , Y ）及び（ X_{END} , Y ）を有する）によって、水平走査領域が得られる。

【0057】粗走査及び精密走査処理は、水平の縁部併合（マージ）フィルタとして作用する。ブロックのサイズが、併合された縁部の間の最大許容距離に関連し、又大きなブロックサイズを処理するための顔面位置検出の速度に直接に影響を与える。粗走査及び精密走査処理によって、頭部の頂部についての候補位置を特定する。この場合、頭部外形に対応する縁部データは、他の対象物に対応するデータによって概して妨げられないことが特徴である。精密走査後、顔面位置検出器96が、画像内の頭部の頂部を含む水平セグメントを特定する。

【0058】楕円適合器102が、座標（ X_{START} , Y ）及び（ X_{END} , Y ）によって定められるラインセグメントを走査する。セグメントの各ポイントにおいて、

種々のサイズ及び形状比を有する楕円が適合性を定められる。この場合、楕円の最頂点は水平走査セグメント上に位置する。よく一致した事例が顔面造作検出システム14のメモリ(図5には図示しない)に設けられたリストにエントリされる。楕円適合器102によるセグメント上での探索が行われた後、顔面位置検出器96が粗相先98を用いて、入力バイナリ縁部データ信号の処理を継続する。

【0059】一実施例においては、「傾斜ゼロ」($\theta = 0$)の楕円が、計算の単純さから入力画像に適合する。与えられた楕円の、データに対する適合性は、楕円輪郭上及び楕円境界上のそれぞれについてバイナリ画素デー

$$I_E(i, j) = \begin{cases} 1 & \text{楕円の内部又は|E|上にあり、} \\ 0 & \text{楕円の外部にあり、} \end{cases} \quad (3)$$

【0061】もし与えられた画素が楕円の内部(又は楕円上)にあり、且つその与えられた画素の近傍($(2L+1) \times (2L+1)$ 画素のサイズを有する)における画素のうちの少なくとも1個がその楕円の内部にない場

$$(i, j) \in C_i \Rightarrow I_E(i, j) = 1$$

及び

$$\sum_{k=i-L}^{i+L} \sum_{l=j-L}^{j+L} I_E(k, l) < (2L+1)^2 \quad (5)$$

【0062】もし与えられた画素画素が楕円の外部にあり、且つその与えられた画素の近傍($(2L+1) \times (2L+1)$ 画素のサイズを有する)における画素のうちの少なくとも1個がその楕円の内部にある場合、すな

$$(i, j) \in C_e \Rightarrow I_E(i, j) = 0$$

及び

$$\sum_{k=i-L}^{i+L} \sum_{l=j-L}^{j+L} I_E(k, l) > 0 \quad (7)$$

【0063】パラメータLは、楕円の輪郭及び境界の望まれる厚さを定める。Lは例えば、1又は2画素にセットされる。このような輪郭及び境界の画素に対して、正

$$I_i = \frac{1}{|C_i|} \sum_{(m, n) \in C_i} w_{m, n} p(m, n) \quad (8)$$

及び

$$I_e = \frac{1}{|C_e|} \sum_{(m, n) \in C_e} w_{m, n} p(m, n) \quad (9)$$

ここに、 $p(m, n)$ はバイナリ画像データを表し、 $|C_i|$ 及び $|C_e|$ はそれぞれ C_i 及び C_e の集合の要素数(カーディナル数)を表す。 $w_{m, n}$ は、図3に示すよう

タの正規化加重平均強さ I_i 及び I_e を計算することによって定められる。楕円輪郭及び楕円境界は、そのノンパラメトリック形式によってよく定義されるけれども、画像データのラスタ化(空間サンプリング)には、実際の画像画素への連続楕円曲線のマッピング(写像)が必要となる。

【0060】本発明の顔面造作検出システム14によって行われる楕円適合処理に対する楕円曲線は、下で述べるように定められる離散曲線である。 $I_E(i, j)$ を楕円Eの内部又は楕円E上にある点の集合(セット)についての特性関数とすると、次式が成立する。

【数2】

合、すなわち次の式(4)及び式(5)が成立する場合に、その与えられた画素は、その楕円輪郭上にあるとして分類される。

【数3】

(4)

わち次の式(6)及び式(7)が成立する場合に、その与えられた画素は、その楕円境界上にあるとして分類される。

【数4】

(6)

規化加重平均強さ I_e 及び I_i は、次の式(8)及び式(9)のように定義される。

【数5】

な、楕円の適合がより確かな領域である「楕円の上部1/4」におけるデータの寄与を強化するための加重ファクタである。

【0064】すなわち、 $w_{m,n}$ は、次の式(10)で表される。

$$w_{(m,n)} = \begin{cases} w > 1 & (\text{もし } (i,j) \in Q_u \text{ ならば} \\ 1 & (\text{もし } (i,j) \notin Q_u \text{ ならば} \end{cases} \quad (10)$$

ここに、 Q_u は、図3に示す「楕円の上部1/4」を示す。

【0065】本実施例においては、重み $w=1.5$ が用いられる。楕円の輪郭及び境界の「長さ」に関する正規化は、異なるサイズの楕円に対応するためにも行われる。概して、楕円は、 I_i の値が最大値 $I_{MAX}=(3+w)/4$ に近いような高い値である場合、及び I_i の値がゼロ値に近いような低い値である場合に、楕円形状のデータに適合する。このような最大化最小化合同条件が、次の式(11)でモデル適合比 R_m を定義することにより、単一量の最大化に変換される。

【0066】

【数7】

$$R_m = \frac{1+I_i}{1+I_e} \quad (11)$$

この場合、 R_m の値が高いほど、入力画像内の頭部外形に、候補の楕円がよりよく適合することになる。例えば完全に楕円形状のデータは、 $I_i=I_{MAX}$ 、 $I_e=0$ 、及び $R_m=1+I_{MAX}$ に対応するデータに最もよく合致した楕円である。

【0067】楕円適合器102は、条件「 $I_i > I_{imin}$ 」及び「 $I_e < I_{emax}$ 」を満足する楕円を適合させることにより、不的確な候補をフィルタ処理して排除する。ここに、 I_{imin} 及び I_{emax} は予め定められたパラメータである。モデル適合比 R_m は、パラメータ I_i 及び I_e の絶対値に対してよりもこれらのパラメータの相対値に対して、より感応度が高い。

【0068】いくつかのビデオ画像においては、顔面を囲む区域(肩部を含む)の部分的な重なりによる閉塞及び顔面を囲む区域(肩部を含む)の動きに起因して、楕円の或る弧の部分だけしか区別できない場合がある。上記のしきい値及びモデル適合比 R_m を用いることによって、顔面位置検出器96が、このような弧の部分を見出しこれに狙いを付けて追跡して、ひどい閉塞状態の顔面であってもその位置を求めることができる。

【0069】顔面位置検出器96は、適合状態のよい楕円を1個以上検出するので、信頼しきい値 ΔR_{min} 及び ΔI_{emin} を用いて、除外プロセスを実行して最終候補を選択する。もしよく適合する楕円についての R_m の値が、第2のよく適合する楕円についての R_m の値よりも ΔR_{min} を超える差で高い場合、最初の楕円が選択される。

【0070】そうでなければ、もし2個の楕円間の境界強さの差が ΔI_{emin} よりも大きい場合、より小さい

【数6】

I_e 値の楕円が選択される。又もし2個の楕円間の境界強さの差が ΔI_{emin} よりも小さい場合、より大きい R_m 値の楕円が選択される。

【0071】よく適合する楕円によって顔面外形を定めた後、顔面位置検出器96は、そのよく適合する楕円のパラメータから顔面位置信号を生成して、符号化制御器16(図1～図2)に供給する。符号化制御器16は、顔面位置信号を用いて、顔面位置に対応する画像内の区域の量子化を増大させる。

【0072】顔面位置検出器96を用いての顔面外形の位置探索に加えて、眼鼻口部領域検出器108を用いた顔面位置探索も行われる。この場には合、図3に示すような楕円領域が、図4に示すような楕円の矩形ウィンドウ部とその補足部分すなわち楕円の残部とに分割される。眼鼻口部領域検出器108が、顔面位置検出器96から検出された顔面外形の楕円パラメータを受信して処理を行って、矩形ウィンドウが眼及び口に対応する顔面の領域を捉えるように、矩形ウィンドウを位置させる。

【0073】眼鼻口部領域検出器108が、文献(F. Lavagetto et al., "Object-Oriented Scene Modeling for Interpersonal Video Communication at Very Low Bit Rate," SIGNAL PROCESSING: IMAGE COMMUNICATION, VOL. 6, 1994, pp. 379-395)に述べられている基本手順を用いて眼/口領域を特定する。

【0074】眼鼻口部領域検出器108は又、被写体がカメラに直接に面しておらず、被写体の顔面に毛髪が存在しそして/又は被写体が眼鏡をかけており、被写体が白色人種の皮膚色素形成がない場合の、入力ビデオ画像内の眼鼻口部領域の検出も行う。眼鼻口部領域検出器108は、鼻を通り口と交差する長手軸に関しての顔面造作の一般的な対称状態を利用して、眼鼻口部領域の検出を確固たるものに行っている。この場合、対称軸は、画像の垂直軸に関して傾斜している。

【0075】眼鼻口部領域の検出は又、ビデオを用いたテレビ会議の際に生じるような、被写体がカメラを直接見ない場合にも行われる。

【0076】眼鼻口部領域検出器108が、探索領域特定器110を用いて探索領域を定める。この場合、眼鼻口部ウィンドウの位置決め推定値を得るために、楕円顔面外形の中心(x_0 , y_0)が用いられる。眼鼻口部ウィンドウの中心についての探索領域は、 $S \times S$ 画素のサイズの正方形の領域である。なお実施例においては $S=12$ である。図4に示すように、眼鼻口部ウィンドウは、顔面外形の短軸及び長軸に関連する固定サイズ $W_w \times h$ を有するように選択される。

【0077】それから眼鼻口部領域検出器108が、探索領域走査器112を用いて探索領域に連関するデータを処理する。この場合、探索量域内のウインドウ中心の各項補位置 (x_k, y_k) について、対称値又は関数値が、顔面軸に関して定められる。顔面軸が、ウインドウの中心の回りに離散角度値だけ回される。一実施例にお

$$S(x_k, y_k, \theta_k) = \frac{1}{A(R)} \left(\sum_{(m,n) \in R \cap W_u} w a_{m,n} + \sum_{(m,n) \in R \setminus W_u} a_{m,n} \right) \quad (12)$$

ここに、 $A(R)$ は、図4に示す台形 R のカーディナル数、すなわち画素面積、 $R \setminus W_u$ は R と W_u の集合差である。

$$a_{m,n} = \begin{cases} 1 & (\text{if } p(m,n) = p(S(m,n)) = 1 \text{ ならば}) \\ \frac{1}{2} & (\text{if } p(m,n) = p(S(m,n)) = 0 \text{ ならば}) \\ 0 & (\text{他の場合}) \end{cases} \quad (13)$$

そして w は1より大きい加重ファクタである。 w の値は、 W_u におけるデータが式(12)の対称値に顕著に寄与するように定められる。矩形ウインドウを W_u 及び R の領域に分割することによって、眼、鼻、及び口に概略対応するデータがウインドウの位置決めに適用されることとなり、又この位置決めは、ほぼ対称的な領域である眼部のデータに依存することとなる。

【0080】眼鼻口部領域検出器108は又、最小密度 D_{\min} より低いデータポイント密度を有するウインドウとして定義された不適格な候補を除去する。それから眼鼻口部領域検出器108が、結果として得られる台形領域 R のパラメータに対応する眼鼻口部領域信号を生成する。この眼鼻口部領域信号は、符号化制御器16によって用いられ、画像内の顔面の眼、鼻、及び口に対応する画像内の台形領域 R 内の画像データの量子化が精細化(リファイン)され、改善される。

【0081】顔面造作検出システム14からの顔面位置信号と眼鼻口部領域信号とが、CCITT勧告H.261標準、参照モデル8(RM8)を実現した符号化制御器16に供給される。CCITT勧告H.261標準は、全てのAC係数に対するデッドゾーンと同一の一樣量子化器を用いたDCT係数の量子化と、8画素のステップサイズを有するDCT係数に対する8ビットの一樣量子化とを定めている。したがって、知覚的周波数加重はない。

【0082】AC係数量子化器のステップサイズは、パラメータ Q_p すなわちMQUANTの値の2倍として定められる。この場合、パラメータ Q_p は、上記標準において引用され、マクロブロック(MB)のレベルまで表示される。本説明においては、マクロブロックを「MB」とも略称する。11×3MBの矩形配列(アレイ)をブロックグループ(GB)と定義する。本実施例に

いては、傾斜値 θ_k は離散値 -10° 、 -5° 、 0° 、 5° 、及び 10° のうちのどれかである。

【0078】対称軸 $B_s((x_k, y_k), \theta_k)$ に関して (m, n) と対称な点としての $S(m, n)$ について、対称値は次の式(12)の用に定められる。

【数8】

$$S(x_k, y_k, \theta_k) = \frac{1}{A(R)} \left(\sum_{(m,n) \in R \cap W_u} w a_{m,n} + \sum_{(m,n) \in R \setminus W_u} a_{m,n} \right) \quad (12)$$

【0079】又 $a_{m,n}$ は次の式(13)によって定められる。

【数9】

$$a_{m,n} = \begin{cases} 1 & (\text{if } p(m,n) = p(S(m,n)) = 1 \text{ ならば}) \\ \frac{1}{2} & (\text{if } p(m,n) = p(S(m,n)) = 0 \text{ ならば}) \\ 0 & (\text{他の場合}) \end{cases} \quad (13)$$

において受信され処理されたビデオ画像は、360×240画素の解像度を有し、結果として画面(フレーム)当たり合計10ブロックグループとなる。

【0083】概して、参照モデル8を用いて、十分小さな値を有する一連のDCT係数を除去する「可変しきい値」手法によって、ジグザグ走査されたDCT係数におけるラン長さについての長さの増加が実現される。可変しきい値は、量子化に先立って適用され、概して符号化効率の改善に有効であり、特に比較的低いビットレートにおいて有効である。MC/ノーマCの決定は、予め定められた曲線に基づくマクロブロックの値及び転置されたマクロブロック差に基づく。

【0084】同様に、イントラ/ノンイントラの決定は、本来のマクロブロックと動作補償されたマクロブロックとの変動の比較に基づく。P個の画面内の予測されたマクロブロックは、もし量子化後にそれらの動作ベクトルがゼロであり且つそれらのブロックの全てがゼロ要素を有する場合、スキップされる。出力バッファがオーバーフローする場合にも、スキップされる。

【0085】レート制御が第1の画面(I画面)から開始される。この画面は、16画素の定数 Q_p で符号化される。出力バッファが、50%の占有率にセットされる。残りの画面 Q_p は1つのブロックグループ内のマクロブロックの各ラインの開始時に適応される。したがって Q_p は、各ブロックグループ内で3回適応される。

【0086】各マクロブロックの送信後にバッファ占有率が点検され、もしオーバーフローが生じる場合には、次のマクロブロックがスキップされる。したがって、バッファのオーバーフローは少量且つ一時的なものとなり、オーバーフローを引き起こしたマクロブロックは送信される。

【0087】 Q_p が次の式(14)の関係に基づきバ

ファ占有率について更新される。

【数10】

$$Q_{pi} = \min \left\{ 31, \left\lfloor \frac{B_i}{B_{\max}/32} \right\rfloor + 1 \right\} \quad (14)$$

ここに、 Q_{pi} はマクロブロック*i*について選択された Q_p の値、 B_i はマクロブロックの符号化に先立っての出力バッファ占有率、そして B_{\max} は出力バッファのサイズである。

【0088】ビデオ信号だけについて、与えられたビットレート $q \times 64 \text{ kbps}$ に対して、 $6,400 \times q$ ビットのバッファサイズが用いられる。一実施例においては、 $6,400$ ビットのバッファサイズが用いられる。

【0089】モデルに補助された符号化によって、画像を見る者にとって知覚的に重要な領域のような、画像内の異なる領域において異なる「品質レベル」が割り当てられる。参照モデル8の仕様が用いられる人対人の通信の場合に用いられるような低いビットレートの符号化に対して、マクロブロックが各ブロックグループ内で通常の左から右へ、上から下への順序で符号化され、量子化器の選択が現バッファ占有レベルに基づいて行われる。

【0090】各ビデオ画像についての予め定められたビット予算内に留まり且つ／又はバッファのオーバフローを回避しながら、関心対象の領域により多くのビットを割り付けるために、このようなマクロブロックの符号化においてマクロブロックの位置探索が用いられる。

【0091】したがって、符号化は、残りの画像領域により少なくビットを割り付けるように制御される。符号化は、1つの画像内の関心の対象である*M*個の領域 R_1, R_2, \dots, R_M について、対応する面積 A_1, A_2, \dots, A_M において行われる。この場合、領域間には重複はない。

$$\beta_0 = \frac{\beta A - \sum_{i=1}^M \beta_i A_i}{A - \sum_{i=1}^M A_i}$$

ここに、次の式(17)に基づき、相対的平均品質 $\gamma_i = \beta_i / \beta$ ($i = 0, \dots, M$)。

$$\gamma_0 = \frac{A - \sum_{i=1}^M \gamma_i A_i}{A - \sum_{i=1}^M A_i}$$

ここに、もし全ての $i > 0$ に対して $\gamma_i > 1$ の場合、 $\gamma_0 < 1$ である。

【0095】上記の式(14)を一般化する次の式(1

$$Q_{pi} = f(B_i)$$

関数 $f(\cdot)$ は入力ビデオ信号に依存する。

【0096】出力バッファの動作は、次の式(19)で

$$B_i = B_{i-1} + c(i-1) - t$$

ここに、 B_i はマクロブロック*i*の符号化に先立っての

すなわち、 $i \neq j$ のとき、交差関係にある R_i 及び R_j はゼロ集合である。領域は凸であることを要しない。画像全体を含む矩形領域を R_I と称し、その面積を A と称する。

【0092】一実施例において目標予算レートが B_r でありバッファサイズが B_{\max} である場合、各マクロブロックの符号化には平均して β ビットが用いられる。

【0093】パラメータ $\beta_1, \beta_2, \dots, \beta_M$ は、関心対象の領域の各々の符号化についての、マクロブロック当りの目標平均ビット数を表す。概して、 $\beta_i > \beta$ は、関心対象領域 R_i 内の品質が改善されていることを示す。関心対象の領域のどれにも属さない画像領域を R_0 、対応する面積を A_0 、そしてマクロブロック当りの平均ビット数を β_0 と表示する。与えられた平均ビット予算値を満足させるためには、次の式(15)が成立する。

【数11】

$$\sum_{i=0}^M \beta_i A_i = \beta A \quad (15)$$

【0094】与えられたパラメータ $\beta_1, \beta_2, \dots, \beta_M$ に対して、全ての対象物の外にある画像領域について同等の平均品質が定められ、領域の、望まれる平均符号化品質及びそれら領域のサイズによって定められる。式(15)は、次の式(16)の形でも表現できる。

【数12】

$$(16)$$

【数13】

$$(17)$$

8)に基づいて作動する符号器の一般的なレート制御動作について、

記述される。

$$(18)$$

$$(19)$$

ここに、 B_i はマクロブロック*i*の符号化に先立っての 50 バッファ占有率、 t はマクロブロック当りのビット数で

21

表す平均目標レート、そして $c(i)$ は i 番目のマクロブロック及びその直前のオーバーヘッド情報、例えばヘッダを符号化するために消費されたビット数である。関数 $c(i)$ は入力ビデオ信号及び現 Q_p 値に依存し、現 Q_p 値は又、関数 $f(\cdot)$ の選択に依存する。

【0097】式(18)及び式(19)は下で述べるように変換されて、位置に依存しモデルに補助される動作が得られる。この場合、本発明の符号化制御器16は、関心対象の内部領域であるマクロブロックにより多くのビットを消費し、関心対象の内部領域ではないマクロブロックにより少なくビットを消費するように目標レートを10 変調するためのバッファレート変調器18(図1)を有する。式(19)中のレート t はこれで位置に依存することとなり、次の式(20)で表される。

【0098】

【数14】

$$t_i = \gamma_{\zeta(i)} t$$

(20)

$$B_i = B_{i-1} + c_{\zeta(i-1)}(i-1) - \gamma_{\zeta(i)} t$$

(21)

ここに、消費されたビットの数 $c_{\zeta(i)}$ は領域に依存する。

領域 k におけるバッファの静止挙動について、又式(21)の両辺に対して期待動作を行うと、領域 k についての平均レートは次の式(22)で表される。

【数16】

$$\bar{c}_k = \gamma_k t$$

(22)

【0101】もし γ_i の値が、式(15)によって与えられる予算制約を満足する場合、全平均レートはレート t である。一定のレート t で内部を空にする通常の非変調出力バッファで作動するシステムについて、バッファのオーバーフロー又はアンダフローが回避されるように式(19)及び式(21)が追跡される。

【0102】式(21)を満足する変調された「仮想」バッファが用いられて、式(18)の関数 $f(\cdot)$ を介して Q_p の生成が駆動される一方、実際のバッファがモニタされて、オーバーフローの場合にマクロブロックをス

$$Q_{pi} = g(B_i, i) = f\left(\frac{B_i}{\mu_{\zeta(i)}}\right)$$

(23)

ここに、 μ_i は画像の各領域についての変調ファクタである。 40

【0105】バッファサイズ変調器20は式(23)を実現して、「 $\gamma_i < 1$ 」及び「 $\mu_i < 1$ 」である関心の低い領域で作動してバッファ占有率が実際よりも高いことを表示し、又「 $\gamma_i > 1$ 」及び「 $\mu_i > 1$ 」である顔面外形及び眼鼻口部領域のような関心の高い領域で作動してバッファ占有率が実際よりも低いことを表示する。したがって、画像内のマクロブロックの位置が顔面領域と一致するかどうかにより、 Q_{pi} 値がより高い値又はより低い値へ押しやられる。

22

この場合、特性関数 $\zeta(i)$ がマクロブロック i の位置をマクロブロック i の属する領域に連関させる。そしてもしマクロブロックがその画素の少なくとも1個がその特定の領域の内部にある場合に、そのマクロブロックはその領域に属するものとみなされる。

【0099】したがって、顔面造作検出システム14からの顔面位置信号及び眼鼻口部領域信号を用いて、バッファレート変調器18がパラメータ γ を、領域特性関数に連関する画像内の顔面領域においてパラメータ γ が1よりも大きくなるように生成する。それからバッファレート変調器18が、式(20)を実現して、検出された顔面外形と眼鼻口部の造作とに対応する画像内の領域の符号化レートを増加させる。

【0100】バッファの動作は今、次の式(21)で記述される。

【数15】

キップさせる。仮想バッファがオーバーフロー下場合、対応動作は取られず、 $f(\cdot)$ に依るが一般に Q_p が最大値を割り当てられる。

【0103】CCITT勧告H.261によって用いられる走査手法のような、マクロブロック走査において、1個の領域のマクロブロックの連続ランに含まれるマクロブロックは僅か1個又は2個であり、結果として、全体のビット分布が非対称となる。この場合、1個の領域の最も左のマクロブロックは、最も右のマクロブロックに比して比較的高い Q_p を有する。

【0104】本発明の符号化制御器16は、バッファサイズ変調器20を実現して、バッファサイズ変調を行う。式(18)が、次の式(23)のように変形される。

【数17】

【0106】詳しくは、高い符号化品質領域をより低い符号化品質領域から符号化する場合に、バッファ占有率は低い。例えば、外部領域においてバッファ占有率は平均で B_{\max}/γ_0 よりも低い。したがって、概して、高い符号化品質領域内でブロックを符号化しながら、生成されたビット数の急速な増加を吸収するに十分なバッファスペースが利用可能となる。

【0107】更に、マクロブロック走査パターンに起因して或る領域の一連のマクロブロックが別の領域の一連のマクロブロックと交替し、したがって「救援」インタ 50 バル(時間間隔)が存在することになり、この間に出力

バッファは内容の排出（ドレイン）が可能となる。

て次の式（24）が得られる。

【0108】それから式（23）を式（14）に適用し

【数18】

$$Q_{p_i} = \min \left\{ 31, \left\lceil \frac{\frac{B_i}{\mu_{\zeta(i)}}}{R_{\max}/32} \right\rceil + 1 \right\} \quad (24)$$

【0109】参照モデル8（RM8）が、各ブロックグループにおいてマクロブロックの各ラインの開始時に Q_p を更新し、バッファサイズ変調器20が、「 $\gamma_i > 1$ 」の領域の内部にある各マクロブロックについて Q_p を更新させる。したがって、バッファレート変調がレート制御動作をして、指定された数のビットを関心対象の領域内で付加的に消費させ、一方バッファサイズ変調が、各領域のマクロブロック内でのこれらの付加的ビットの一樣分布を確実にする。

【0110】なお、本発明のバッファレート変調及びバッファサイズ変調の手法は、動作表示器等を勘案したレート制御方式を含んで一般にどのようなレート制御方式にも適用されることを理解されたい。

【0111】以上、本発明の顔面造作検出システム及び方法を特にいくつかの実施例に関して述べたが、本発明の種々の変形例が可能で、それらはいずれも本発明の技術的範囲に包含されることが、この技術分野の当業者であれば理解されよう。

【0112】

【発明の効果】以上述べたごとく、本発明によれば、CITT勧告H.261標準対応の低ビットレートでのビデオ符号化において、被写体人物の顔面外形及び顔面造作位置を近似モデルを用いて自動的に検出でき、顔面及び顔面造作の画像を適切な符号化品質で採取送信することが可能となる。したがって、テレビ会議等の画面上での対人の相互通信の品質及び効率が改善される。又、読唇に頼る聴覚障害者の場合のような或る種の状況において、本発明により顔面の造作の良好な表現が得られるので、障害者にも有利である。

【図面の簡単な説明】

【図1】符号化復号化器（コーデック）のブロック図である。

【図2】ソースコードのブロック図である。

【図3】楕円及び連関するパラメータの説明図である。

【図4】矩形領域及び連関するパラメータの説明図である。

【図5】本発明に基づく顔面造作検出器のブロック図である。

【図6】前処理プロセッサのブロック図である。

【符号の説明】

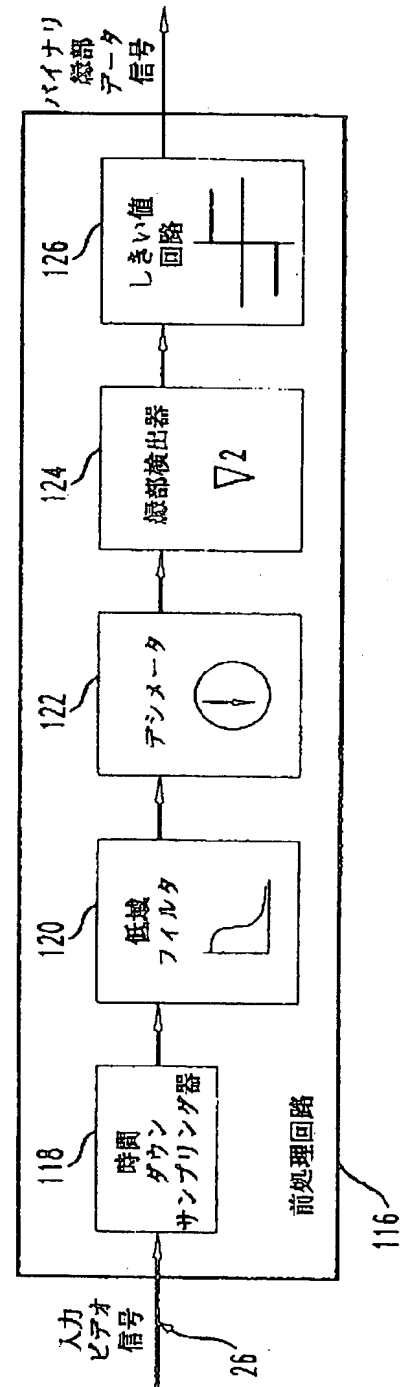
- 10 符号化復号化器（コーデック）
- 12 対象物位置探索器
- 14 顔面造作検出システム
- 16 符号化制御器

- 18 バッファレート変調器
- 20 バッファサイズ変調器
- 22 ビデオ符号器
- 24 ビデオ復号器
- 26 入力ビデオ信号
- 28 外部制御信号
- 30 符号化されたビットストリーム（出力）
- 32 ソース符号器
- 34 ビデオ多重符号器
- 36 送信バッファ
- 38 送信符号器
- 40 符号化されたビットストリーム（入力）
- 42 受信復号器
- 44 受信バッファ
- 46 ビデオ多重復号器
- 48 ソース復号器
- 50 出力ビデオ信号
- 52 減算器
- 54、66 切換器
- 56 変形ユニット
- 58 量子化器
- 60 逆量子化器
- 62 逆変形ユニット
- 64 加算器
- 68 ループフィルタ
- 70 動作補償遅れ付き画像メモリ
- 72 楕円E
- 74 中心
- 76 半長径A
- 78 半短径B
- 80 傾斜角 θ_0 （楕円長軸の）
- 82 矩形領域
- 84 顔面外形の対称軸（矩形領域における） A_s
- 86 画像垂直線
- 88 ウインドウ
- 90 ウインドウ中心
- 92 角度 θ_i （台形側線の）
- 94 顔面位置前処理プロセッサ
- 96 顔面位置検出器
- 98 粗走査器（スキャナ）
- 100 精密走査器
- 102 楕円適合器
- 104 顔面位置信号
- 106 眼鼻口部領域前処理プロセッサ

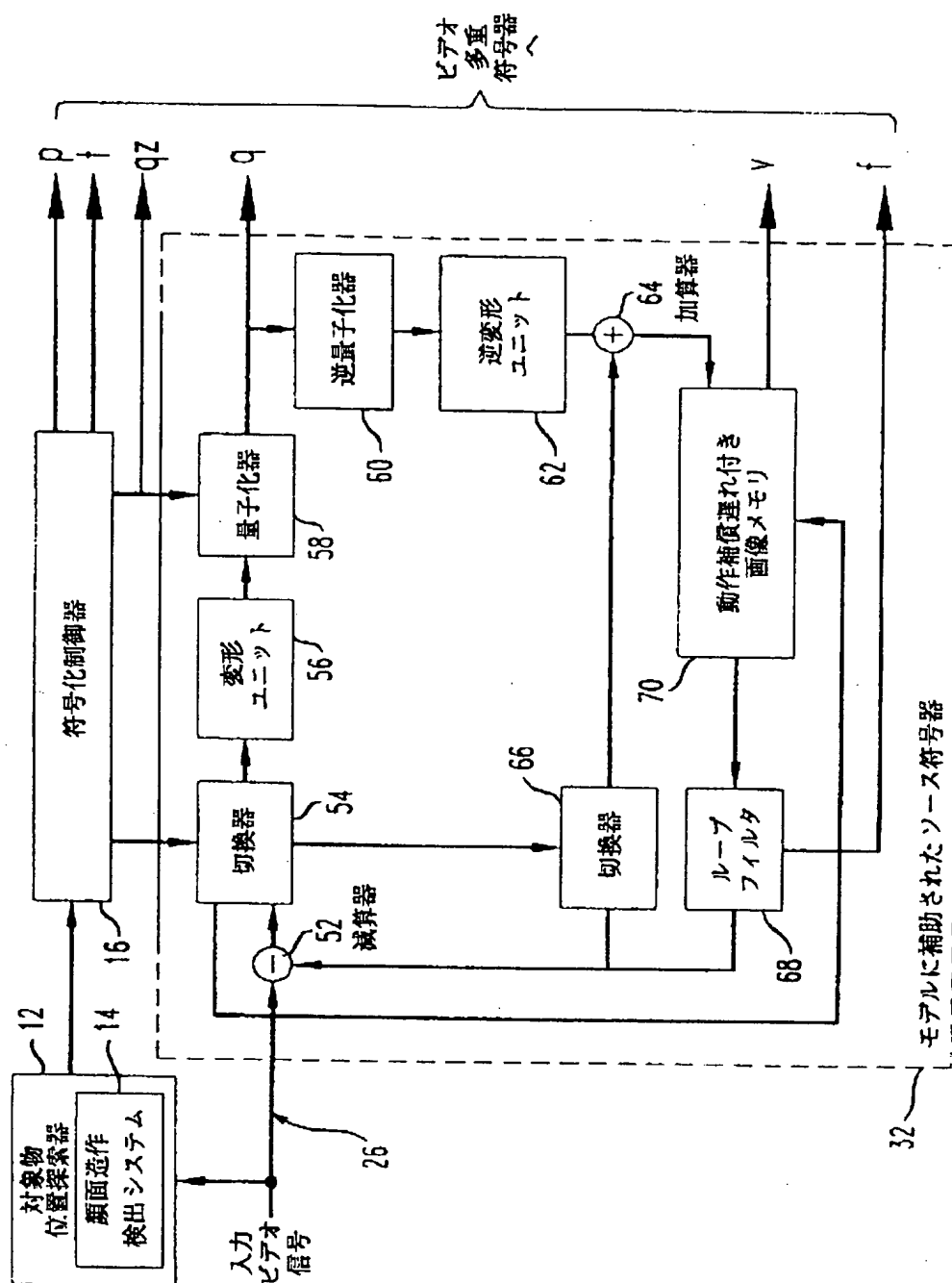
26

- 1 1 8 時間ダウンサンプリング器
1 2 0 低域フィルタ
1 2 2 デシメータ (1 / 1 0 減数器)
1 2 4 縁部検出器
1 2 6 しきい値回路

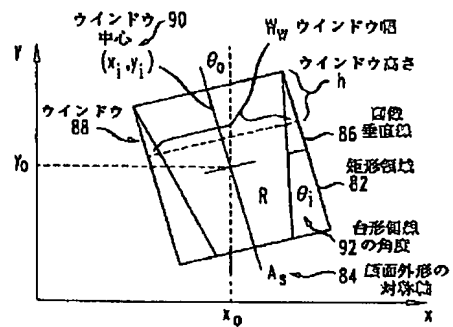
【图 6】



ビデオ多重符号器へ



【图 4】



図面動作検出システム

入力ビデオ信号 (26)

94 図面位置前処理プロセッサ

96 図面位置検出器

98 位置寄存器

100 位置定数器

102 箱門辺合器

104 図面位置信号

106 図面口部領域前処理プロセッサ

108 図面口部領域検出器

110 探索領域特定器

112 探索領域走査器

114 図面口部領域信号

14

符号44参照

(72)発明者 アルナウド エリック ジャックイン
アメリカ合衆国, 10014 ニューヨーク、
ニューヨーク、パンク ストリート 33、
アパートメント 21